ON PERIODIC ASYMMETRIC EXTRAPOLATION

TETIANA BOIKO, OLEG KARPENKOV, BULAT RAKHIMBERDIEV

Abstract. In this paper we develop a new techniques for asymmetric approximation of discrete functions originated in seasonal customer demand extrapolation. We adapt the techniques for two different settings: so called pull and push models. Our main goal here is to find effectively the loss minimizing extrapolations. For both models we discuss several features concerning sampling, approximation, and extrapolation.

1. Introduction

In this paper we develop a techniques for approximation of discrete functions that has a "nearly" periodic behaviour. The problem has an applied origins, it is closely related to forecasting (i.e., to extrapolation) of recurrent patterns in consumer demand in superstores. There are two different models of store operation which we deal with, they are as follows.

- *Seasonal pull model:* In this model the store immediately replenish a necessary amount of items for the next unit of time.
- *Seasonal push model:* In this model the store order items several units of time in advance. If one uses pull extrapolation in such settings then storage loss expectation will be higher.

While developing extrapolation models we take into account the following two types of losses.

- *Out-of-stock loss*: here the company looses customers due to shortage of items in the storage.
- *Storage loss*: storing exceeding number of items is also expensive. For instance this is essential for the market of electronics when the demand of old models strongly decreases when new models are offered for sale.

In both cases of seasonal pull and push models we aim to find the extrapolation that minimizes the loss functional. In this paper we develop a method to find one of the minimizers of the total loss functional explicitly. We show that the set of values of such an extrapolation is a subset of a certain finite set. Then this minimizer can be found by brute force case study (see Theorem 7.1). This method is much faster than standard gradient algorithms for small values of k.

Date: 30 June 2018.
Key words and phrases. periodic functions, approximation, pull and push models, maximizers.

Main stages of seasonal storage push and pull extrapolation. Both push and pull models follow the same general pattern.

Given an observed history of trading we complete the following stages.

- **Sampling and normalization**: On that stage we synthesize a certain set of periodic functions from the input data, in addition we normalize every such function such that the sum over the period equals 1. We call such functions the *normalized periodic particles*. (for an alternative way of sampling, e.g., see [3]).
- **Periodic approximation**: First of all for both seasonal pull and push models we introduce an extrapolation loss functional on the spaces of all normalized periodic particles, constructed above. This functional essentially depends on the set of normalized periodic particles. Further, using the result of Theorem 7.1, we find best fitting periodic approximations that minimize such functionals.
- **Particle discrepancy (weighting)**: Once a best fitting periodic approximation is computed, we can introduce the weigh function on the set of normalized periodic particle. A further normalized periodic particles from the best fitting periodic approximation is, a smaller weight it has (for a general theory of denoising we refer to [6]).
- **Extrapolation minimizers**: So we have constructed the set of normalized particles equipped with denoising weights. Now for both seasonal pull and push models we introduce a weighted extrapolation loss functional on the spaces of all normalized periodic particles. Again, using Theorem 7.1, we find an extrapolation that minimizes these functionals.
- **Rescaling**: The obtained minimizers are computed for normalized periodic particles. So we rescale back to get a true forecast (see Remark 5.7).

Recall several approaches that are in actual use and have proved to be successful. The classical approach to time series forecasting derives from regression analysis. The standard regression model involves specifying a linear parametric relationship between a set of explanatory variables and the dependent variables of the model. The parameters of the model can be estimated in a variety of ways, for example, by least squares method introduced by Gauss in 1794 or more "modern" approaches introduced by N. Wiener (see [9]) and A. Kolmogorov (see [8]).

Methods such as seasonal decomposition, Box-Jenkins and ARIMA are designed to extract seasonal and other cyclical component signals from a series by means of an iterated finite moving average procedure (see [1], [4], and for a general overview [5]). C.C. Holt [7] and P.R. Winters [10] further generalized this method to include a linear component in the extrapolation function. A few years later Brown [2] reformulated the problem in terms of a discounted least squares regression (linear exponential smoothing).

The extrapolation techniques we develop in this paper are within *seasonal time series methods*. The main difference between the listed time series methods and our methods are as follows. First of all we do not assume the error term to be a random variable, instead of that we deal with functional space of all samples optimizing a certain functional on it. Secondly we introduce a special weighted system (which is distinct to the mean absolute percentage error) in order to find the best fitting periodic approximation, which improves the original data. The proposed techniques were successfully implemented in a chain of stores selling consumer electronics.

This paper is organized as follows. In Section 2 we describe the input data and describe sampling we employ. Further in Section 3 we define best seasonal pull extrapolation. In Section 4 we introduce the push extrapolation which generalizes the pull extrapolation. We introduce a denoising weight system in order to reduce noisy nonseasonal effects on the space of normalized periodic particles in Section 5. In Section 6 we show how to find minimizers of periodic σ-discrepancy which is the key point to construct a best weighted seasonal pull extrapolation. Further in this section we discuss criteria of uniqueness of best weighted seasonal pull extrapolations. We conclude this paper in Section 7 with techniques to find explicitly the minimizers in the push model.

2. Input data, sampling, and periodic seasonal approximation

In this section we give basic notions and definitions. In Subsection 2.1 we set the input data for the models we study in this articles. Further in Subsection 2.2 we show how to construct normalized periodic particles. They are used as samples to construct the best approximation. Finally in Subsection 2.3 we define the space of P-periodic seasonal approximations.

2.1. Input data.

The seasonal extrapolations described in this article are computed basing on the following *input data*:

- T: a *total number of observations* (T is a positive integer);
- P: a *seasonal period* is a number of observations which we consider as a period (P is a positive integer such that $P \leq T$);
- p_I: an *out-of-stock loss* value is the price that we pay for one out-of-stock loss ($p_I \geq 0$);
- p_{II}: a *storage loss* value is the price that we pay to store one item per one time unit ($p_{II} \geq 0$);
- $f : \{1, 2, \ldots, T\} \to \mathbb{R}_{\geq 0}$: an *observation data* is a function whose value $f(t)$ is the number of items sold between observations t and $t + 1$ (for $t = 1, \ldots, T$). Additionally we require that f does not have P consequent zero values.

Denote this input data by (T, P, p_1, p_2, f).

Example 1, part 1 of 2. The input data of this example is a sample of real-life data from one company which is a chain of stores selling home electronics. The observation function f is as on the following diagram:

It contains 79 observations, and hence $T = 79$. There were recorded on a monthly basis, i.e. $P = 12$. The out-of-stock and storage loss values were set by the company as $p_I = p_{II} = 1$.

Remark 2.1. While studying real-life examples of observation functions it is clear most of them have strong seasonality behaviour, like with one described in Example 1. Nevertheless the pikes might occur at the neighbouring time segments in different years (as discussed in Example 2 part 1 below), which makes them far from being periodic.

Throughout this paper we will go through the following simpler example.

Example 2, part 1 of 5. The observation function is

$$f = (1, 1, 1, 1, 10; 1, 1, 1, 10, 1; 1, 1, 1, 1, 10; 1, 1, 1, 1, 10).$$

(Here and below we write functions as sequences of values, assuming that the first element of the sequence is the value at time 1, unless otherwise stated.) The observation function is defined for $t = 1, \ldots, 20$ (hence $T = 20$). The seasonal period P is considered to be 5. Finally set $p_I = 3$ and $p_{II} = 1$. So the input data is

$$\Big(20, 5, 3, 1, (1, 1, 1, 1, 10; 1, 1, 1, 10, 1; 1, 1, 1, 1, 10; 1, 1, 1, 1, 10)\Big).$$

Our task is to *compute a best weighted seasonal* $[20, 26]$*-pull extrapolation* (*i.e., a best weighted seasonal pull extrapolation at time 20 till time 26*).

In this example we observe a seasonal behavior, although the observation function is not periodic in classical sense, we have

$$\max_{t\in\{1,\dots,15\}} \big(|f(t{+}5) - f(t)|\big) = |f(10) - f(5)| = 9,$$

which is comparable to the maximum of the function itself (which is 10 here).

2.2. **Sampling: normalized periodic particles.** There are several sampling strategies for construction of approximations. In both pull and push models we use the following natural sampling.

Definition 2.2. Given a triple (T, P, f). For every integer i such that $1 \le i \le T{-}P{+}1$ consider a periodic function $f_i : \mathbb{Z} \to [0, 1]$ with period P defined as follows:

$$f_i(t) = \frac{f(t)}{S_i}, \quad \text{for } t = i, \dots, i + P - 1,$$

where

$$S_i = \sum_{k=i}^{i+T-1} f(k).$$

At all other values of the argument the function f_i is defined by periodicity: $f_i(t + P) = f_i(t)$. The function f_i is said to be a *normalized periodic particle* of a triple (T, P, f). (We assume that f is not identical to zero at any consequent P integers.)

Example 2, part 2 of 5. In our testing example we will have the following normalized periodic particle types of normalized periodic particles:

No.	**Period**	**Graphs**	**Amount**	**Particles**
I	$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{5}{7}\right)$		10	$f_1, \dots, f_4, f_{11}, \dots, f_{16}$
II	$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{5}{7}, \frac{1}{14}\right)$		4	f_6, f_7, f_8, f_9
III	$\left(\frac{1}{23}, \frac{1}{23}, \frac{1}{23}, \frac{10}{23}, \frac{10}{23}\right)$		1	f_5
IV	$\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$		1	f_{10}

2.3. P**-periodic seasonal approximations.** In this subsection we describe techniques of seasonal storage loss extrapolation. Later in Section 5 we improve it by introducing denoising weighted for normalized periodic particles. We start with some general notation.

For integers a, b satisfying $a > b$ and $h : \mathbb{Z} \to \mathbb{R}$ we formally set (similar to integration)

$$\sum_{k=a}^{b} h(t) = -\sum_{k=b+1}^{a-1} h(t).$$

Note that as a consequence we have $\sum\limits_{k=a}^{a-1} h(t) = -\sum\limits_{k=a}^{a-1} h(t)$, and hence $\sum\limits_{k=a}^{a-1} h(t) = 0$.

Definition 2.3. Given $h : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ and an integer t_0. A *discrete distribution function* for h with respect to t_0 is the function

$$H^{t_0} : \mathbb{Z} \to \mathbb{R}, \quad \text{where} \quad H^{t_0}(t) = \sum_{k=t_0}^{t} h(k).$$

Let us continue with the following formal definition.

Definition 2.4. A periodic function $g(t) : \mathbb{Z} \to [0, 1]$ with period P is said to be a *P-periodic seasonal approximation.* The space of P-periodic seasonal approximation is naturally associated with $[0, 1]^P$.

3. Pull extrapolation model

In this section we describe the pull extrapolation model. First we start with the notion of Macaulay brackets that describe the ramp function:

$$\langle x \rangle = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

Let us define the loss of an extrapolation g with respect to an observation function f.

Definition 3.1. Let p_I and p_{II} be the out-of-stock and storage loss values, and let t_0 be an integer. Consider a pair of periodic functions $f, g : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ with period P where f is not zero at least at one point. A *periodic δ-storage loss* of g with respect to f is

$$\delta_f^{t_0}(g) = p_I \cdot \big\langle f(t_0) - g(t_0) \big\rangle + p_{II} \cdot \Big(\sum_{t=t_0}^{+\infty} \big\langle g(t_0) - F^{t_0}(t) \big\rangle \Big). \tag{1}$$

where F^{t_0} is the discrete distribution function for f with respect to t_0 (recall that $F^{t_0}(t_0) = f(t_0)$).

Remark 3.2. Here we consider $\delta_f^{t_0}(g)$ as a single brick to construct best P-periodic seasonal approximations. Note that $\delta_f^{t_0}(g)$ depends entirely on $g(t_0)$, its behavior therefore is similar to the behavior of generalized Dirac δ-functions in continuous settings.

For a P-periodic function h we set

$$E_P(h) = \sum_{t=1}^{P} h(t). \tag{2}$$

Definition 3.3. Given an input data (T, P, p_1, p_2, f) and a P-periodic seasonal approximation g.

- *Periodic δ-discrepancy* of g with respect to the input data is defined as follows:

$$\Delta_f^{t_0}(g) = \frac{1}{T-P+1} \sum_{i=1}^{T-P+1} \Big(\delta_{f_i}^{t_0}(g) \cdot E_P(f_i)\Big). \tag{3}$$

- A *periodic discrepancy* of g with respect to f is

$$\mathcal{D}_f(g) = \sum_{t=1}^{P} \Delta_f^t(g). \tag{4}$$

- Global minimizers of $\mathcal{D}_f$ are said to be *best P-periodic seasonal approximations.*
- The discrepancy for a best P-periodic seasonal approximation is called the *measure of periodicity/sesonality* of the observation function f.

Example 2, part 3 of 5. For our example the best 5-periodic seasonal approximation for the normalized periodic particles is the following function:

$$\Big(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{10}{23}, \frac{5}{7}\Big).$$

Definition 3.4. Given an input data (T, P, p_1, p_2, f), integers $t_0 \leq t_1$ and an interval (finite or infinite) I such that $[t_0, t_1] \subset I$. Consider a function $g : I \cap \mathbb{N} \to [0, 1]$.

- A *pull extrapolation loss* at time t_0 till time t_1 is given by the following expression

$$\Omega_f^{t_0,t_1}(g) = \sum_{t=t_0}^{t_1} \Delta_f^t(g). \tag{5}$$

- *Best seasonal pull extrapolation* at time t_0 till t_1 is a global minimizer of the functional $\Omega_f^{t_0,t_1}(g)$.

We conclude this section with two general remarks.

Remark 3.5. Note that best P-periodic seasonal approximations do not necessarily sum up to 1 at the period. For instance this is the case in Example 2 (see Example 2 part 3). The reason for that is as follows: in order to catch the customers it is worthy to store items in some excess.

Remark 3.6. It is interesting to observe a continuous analog of a best P-periodic seasonal approximation. In this case all functions are defined on intervals. For *continuous periodic δ-discrepancy* we have:

$$\tilde{\Delta}_f^{t_0}(g) = \frac{1}{T-P}\int\limits_0^{T-P}\Bigg(\bigg(p_I\big\langle f_\lambda(t_0)-g(t_0)\big\rangle+p_{II}\Big(\int\limits_{t_0}^{+\infty}\big\langle g(t_0)-f_\lambda^{t_0}(s)\big\rangle ds\Big)\bigg)\cdot\Big(\int\limits_0^P f_\lambda(u)du\Big)\Bigg)d\lambda,$$

and a *continuous periodic discrepancy* is $\int_0^P \tilde{\Delta}_f^t(g)dt$. Here f_λ is a periodic extension of f restricted to the segment $[\lambda,\lambda+P]$ and further normalized (i.e., divided by the value of the integral of f over the segment $[\lambda,\lambda+P]$).

4. Push extrapolation model

In this section we briefly give main definitions for push extrapolation model.

Definition 4.1. Let p_I and p_{II} be the out-of stock and storage loss values. Let also t_0 and t_1 be two integers such that $t_1 \geq t_0$. Consider a function $f:\mathbb{Z}\to\mathbb{R}$ with an unbounded from above distribution function F^{t_0} with respect to t_0. Let $g:\{t_0,\ldots,t_1\}\to\mathbb{R}_{\geq 0}$. Then the $[t_0,t_1]$-*push extrapolation loss* of g with respect to f is

$$\begin{aligned}\Lambda_f^{t_0,t_1}(g) =& \sum_{t=t_0}^{t_1}\Big(p_I\cdot\big\langle R_{f,g}^{t_0}(t)\big\rangle + p_{II}\cdot\big\langle -R_{f,g}^{t_0}(t)\big\rangle\Big)+\\ & p_{II}\cdot\sum_{t=t_1+1}^{+\infty}\big\langle R_{f,g}^{t_0}(t_1)-F^{t_1}(t)\big\rangle.\end{aligned}\tag{6}$$

Here the reminder function is defined iteratively

$$\begin{aligned}&R_{f,g}^{t_0}(t_0) = g(t_0)-f(t_0);\\ &R_{f,g}^{t_0}(t{+}1) = \langle R_{f,g}^{t_0}(t)\rangle + g(t+1)-f(t+1), \quad \text{for } t=t_0{+}1,\ldots,t_1.\end{aligned}$$

In other words

$$R_{f,g}^{t_0}(t) = \langle\cdots\langle\langle g(t_0)-f(t_0)\rangle + g(t_0{+}1)-f(t_0{+}1)\rangle\cdots + g(t{-}1)-f(t{-}1)\rangle + g(t)-f(t)$$

Remark 4.2. The periodic δ-storage loss in seasonal pull and push models are related by a simple formula:

$$\delta_f^{t_0}(g) = \Lambda_f^{t_0,t_0}(g).\tag{7}$$

Let us now give weighted analogs in seasonal push model.

Definition 4.3. Let (T,P,p_I,p_{II},f) be an input data as above, let $t_1\geq t_1$ be nonnegative integers. Let also f_i be normalized periodic particles for $i=1,\ldots,T-P+1$.

- Consider $g:\{t_0,\ldots,t_1\}\to\mathbb{R}_{\geq 0}$. Then the *seasonal* $[t_0,t_1]$-*push extrapolation loss of g at time* t_0 is as follows:

$$\mathcal{L}_f^{t_0,t_1}(g) = \frac{1}{T-P+1}\sum_{i=1}^{T-P+1}\Big(\Lambda_{f_i}^{t_0,t_1}(g)\cdot E_P(f_i)\Big).\tag{8}$$

- *Best seasonal* $[t_0, t_1]$*-push extrapolation* are global minimizers of $\mathcal{L}_f^{t_0,t_1}$.

5. Weighted seasonal extrapolation

In Example 2 we have spotted one serious problem with the methods described above. Some of the normalized periodic particles are noisy, they are rather far from an average normalized periodic particle. In particular a from common sense suggests that the constant normalized periodic particle of type **IV**: $(1/5, 1/5, 1/5, 1/5, 1/5)$ is noisy (see Example 2, part 2). The total sum here is 5, it is not 14 as expected. So it has a noisy contribution to every value of the period. In fact the noise of such normalized periodic particles can be reduced by the denoising techniques described in this section.

We introduce weights to seasonal push and pull extrapolation models in Subsections 5.1 and 5.2 respectively. Further in Subsection 5.3 we discuss a particular normalized periodic particle denoising which we use in our weighted seasonal pull and push extrapolation models.

5.1. **Weighted seasonal pull extrapolation.** Let us give the following general definition.

Definition 5.1. Given an input data (T, P, p_1, p_2, f), integers $t_0 \le t_1$, and a finite or infinite interval I such that $[t_0, t_1] \subset I$. Let $\mu = (\mu_1, \dots, \mu_{T-P+1})$ be a collection of positive numbers. Consider a function $g : I \cap \mathbb{N} \to [0, 1]$.

- *Weighted periodic δ-discrepancy* of g with respect to the input data is defined as follows:

$$\Delta_f^{t_0,\mu}(g) = \sum_{i=1}^{T-P+1} \Big(\mu_i(f)\delta_{f_i}^{t_0}(g) \cdot E_P(f_i)\Big). \tag{9}$$

- A *weighted* $[t_0, t_1]$*-pull extrapolation loss* at time t_0 till time t_1 is given by the following expression

$$\Omega_f^{t_0,t_1,\mu}(g) = \sum_{t=t_0}^{t_1} \Delta_f^{t,\mu}(g). \tag{10}$$

- A *best weighted seasonal* $[t_0, t_1]$*-pull extrapolation* is a global minimizer of the functional $\Omega_f^{t_0,t_1,\mu}(g)$.

Remark 5.2. Note that the standard seasonal pull extrapolation loss is the weighted seasonal pull extrapolation loss with all weights being equal to $\frac{1}{T-P+1}$.

5.2. **Weighted seasonal push extrapolation.** Consider a similar general definition for weighted seasonal push settings.

Definition 5.3. Given an input data (T, P, p_1, p_2, f), integers $t_0 \le t_1$. Let also f_i be normalized periodic particles with distributions functions F_i where $i = 1, \dots, T-P+1$. Consider a collection of positive numbers $\mu = (\mu_1, \dots, \mu_{T-P+1})$. Fix a best P-periodic seasonal approximation $\tilde{f}$.

- Let $g : \{t_0, \ldots, t_1\} \to \mathbb{R}_{\geq 0}$. Then the *weighted seasonal* $[t_0, t_1]$*-push extrapolation loss of* g *at time* t_0 is as follows:
$$\mathcal{L}_f^{t_0,t_1,\mu}(g) = \frac{1}{T-P+1} \sum_{i=1}^{T-P+1} \mu_i(f) \Lambda_{f_i}^{t_0,t_1}(g),$$
- A *best weighted seasonal* $[t_0, t_1]$*-push extrapolation* is a global minimizer of $\mathcal{L}_f^{\mu}$.

5.3. **Normalized periodic particle denoising.** In both pull and push models we use the following natural noise function.

Definition 5.4. Let $\tilde{f}$ be a best P-periodic seasonal approximation. Then the *noise* of a P-periodic seasonal approximation g with respect to $\tilde{f}$ is defined as
$$\Theta_{\tilde{f}}(g) = \frac{1}{1 + (\mathcal{D}_{\tilde{f}}(g))^2}.$$

Remark 5.5. The set of all best P-periodic seasonal approximations is naturally ordered lexicographically with respect to their sequence of values. So one can always pick the smallest best P-periodic seasonal approximation with respect to lexicographical order.

Definition 5.6. Let (T, P, p_I, p_{II}, f) be an input data as above, let t_0, t_1 be nonnegative integers satisfying $t_1 \geq t_0$. Let also f_i be normalized periodic particles for $i = 1, \ldots, T{-}P{+}1$. Consider the smallest best P-periodic seasonal approximation $\tilde{f}$ with respect to the lexicographical order. The *denoising weight* of the normalized periodic particle f_i is the following number
$$\nu_i(f) = \frac{\Theta_{\tilde{f}}(f_i)}{\sum_{j=0}^{T-P+1} \Theta_{\tilde{f}}(f_j)}.$$
for $i = 1, \ldots, T{-}P{+}1$.

Example 2, part 4 of 5. In our testing example the denoising weights are as follows

Function No.	**I**	**II**	**III**	**IV**
Weights ν_i	0.076	0.038	0.057	0.027

So the best weighted seasonal $[21, 25]$-pull extrapolation for one period is
$$\Big(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{5}, \frac{5}{7}\Big).$$
Here we have a correction at time 4:

Standard best approximation	**Best weighted** $[21, 25]$**-pull extrapolation**
$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{10}{23}, \frac{5}{7}\right)$	$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{5}, \frac{5}{7}\right)$

Remark 5.7. Assume that we have computed a best weighted seasonal $[t_0, t_1]$-push extrapolation g. Then one should consider

$$E \cdot g : \{t_0, \dots, t_1\} \to \mathbb{R}_{\geq 0},$$

where E is an extrapolation of the expectation rate for the total sum for all observations in the consequent P steps.

The function E can be computed by iteratively applying the above seasonality techniques to the *averaging function* formed by $E_P(f_i)$, i.e., to

$$\big(E_P(f_1), \dots, E_P(f_{T-P+1})\big).$$

This function has a $T-P+1$ entry and a seasonal period P.

Example 2, part 5 of 5. In our example **best weighted seasonal** $[21, 26]$**-pull extrapolation** is then

$$\Big(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{5}, \frac{5}{7}, \frac{1}{14}\Big).$$

In order to get the final extrapolation we multiply the obtained function by the expectation E of a total amount of customers in during a period P (as mentioned in Remark 5.7). In our case it is 14. Finally we have

$$\Big(1, 1, 1, \frac{14}{5}, 10, 1\Big).$$

Example 1, part 2 of 2. Let us finally conclude Example 1. On the following picture the observed function is filled with grey. The best weighted seasonal pull extrapolation is shown with thick line. We extend the best weighted seasonal pull extrapolation for the past history in order to examine the quality of extrapolation.

6. Properties of the weighted pull extrapolation

In this section we prove main statements on weighted pull extrapolation. In Subsection 6.1 we prove basic properties of weighted seasonal pull extrapolation loss. Further in Subsection 6.2 we give a finite list which contains all the values for some best weighted seasonal pull extrapolation. Finally in Subsection 6.3 we discuss the uniqueness of best seasonal pull extrapolation.

6.1. **Basic properties.** Let us collect basic properties of weighted seasonal $[t_0, t_1]$-pull extrapolation loss in the following proposition.

Proposition 6.1. *For every integers $t_0 < t_1$ we have*

i) Let g be a P-periodic seasonal approximation g. Then $\Omega_f^{t_0,t_0+P-1,\mu} = \mathcal{D}_f^{\mu}(g)$.

ii) Let I be an interval containing t_0 and $t_1{+}P$ and let $g : I \cap \mathbb{N} \to [0,1]$. Then $\Omega_f^{t_0+P,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu}(g)$;

iii) Let I be an interval containing t_0 and $t_1{+}P$ and let $g : I \cap \mathbb{N} \to [0,1]$. Then $\Omega_f^{t_0,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu} + \mathcal{D}_f^{\mu}(g)$.

Proof. Item (i): We have

$$\Omega_f^{t_0,t_0+P,\mu}(g) = \sum_{t=t_0}^{t_0+P-1,\mu} \Delta_f^{t,\mu}(g) = \sum_{t=1}^{P,\mu} \Delta_f^{t,\mu}(g) = \mathcal{D}_f^{\mu}(g).$$

The second equality holds since both f_i (for every admissible i) and g are periodic with period P.

Item (ii) holds since both f_i (for every admissible i) and g are periodic with period P.

Item (iii): From the above two items we have

$$\Omega_f^{t_0,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu}(g) + \Omega_f^{t_1+1,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu}(g) + \Omega_f^{1,P,\mu}(g) = \Omega_f^{t_0,t_1,\mu} + \mathcal{D}_f^{\mu}(g).$$

This concludes the proof. □

6.2. **On values of periodic discrepancy.** The following theorem is one of the central theorems in this article.

Theorem 6.2. *Let (T, P, p_I, p_{II}, f) be an input data, let μ be a collection of positive integers, and let t_0 be an integer. Then there exists a P-periodic seasonal approximation g that fulfills the following two conditions:*

— the approximation g minimizes the weighted periodic δ-discrepancy functional $\Delta_f^{t_0,\mu}$.

— the value $g(t_0)$ is contained in the union of all values of $F_i^{t_0}(t)$ for $i = 1, \dots, T-P+1$ and $t = t_0, \dots, t_0+P-1$.

We start with the following simple statement.

Lemma 6.3. *The image of any best weighted P-periodic seasonal approximation is contained in* $[0, 1]$.

Proof. It is clear that reducing the value to 1 or increasing negative value to 0 will reduce the value of the corresponding weighted periodic δ-discrepancy. □

Proof of Theorem 6.2. The functional $\Delta_f^{t_0,\mu}$ is piecewise linear when we vary $g(t_0)$ and fix all the other values. In addition it is linear outside zeroes of Macaulay brackets involved in $\Delta_f^{t_0,\mu}$, see Equation (9). Zeroes of such Macaulay brackets are either at $f_i(t_0)$ or at $F_i^{t_0}(t)$ for some integer $t > 0$. Since be definition $f_i(t_0) = F_i^{t_0}(t_0)$, every non-linearity point is at $F_i^{t_0}(t)$ for some $t \geq 0$.

Since $\Delta_f^{t_0,\mu}$ is bounded from below by zero and piecewise linear, it has a global minimum on the real line. Since $\Delta_f^{t_0,\mu}$ is piecewise linear, one can choose the global minimum at the non-linearity point. As we have shown above, all such points are contained in the set of all values of $F_i^{t_0}$.

From the definition of normalized periodic particles we have

$$F_i^{t_0}(t_0 + P - 1) = 1$$

for all admissible i. Hence all values of F^{t_0} that are in the segment $[0, 1]$ are attained at points $t_0, \dots, t_0+P-1$ respectively. Now the statement of the theorem follows from Lemma 6.3. □

Corollary 6.4. *Let (T, P, p_I, p_{II}, f) be an input data, let μ be a collection of positive integers, and let $t_0 \leq t_1$ be a pair of integers. Then there exists a best weighted seasonal $[t_0, t_1]$-pull extrapolation g such that for every admissible $\hat{t}$ the value $g(\hat{t})$ is contained in the union of all values of $F_i^{\hat{t}}(t)$ for $i = 1, \dots, T-P+1$ and for $t = \hat{t}, \dots, \hat{t}+P-1$.*

Proof. By definition we have

$$\Omega_f^{t_0,t_1,\mu} = \sum_{t=t_0}^{t_1} \Delta_f^{t,\mu}.$$

For every integer t_2 in the segment $[t_0, t_1]$ the value $\Delta_f^{t,\mu}(g)$ at t_2 depends only on $g(t_2)$ and does not depend on the other values of g in the period. Hence g minimizes $\Omega_f^{t_0,t_1,\mu}$ if and only if it minimizes every summand $\Delta_f^{t,\mu}$ in the sum. This reduces Corollary 6.4 to Theorem 6.2. □

6.3. **On uniqueness of a weighted pull extrapolation.** We conclude this section with the following observation.

Remark 6.5. A best weighted seasonal $[t_0, t_1]$-pull extrapolation is uniquely defined if

$$\Big(\sum_{i=1}^{T-P+1} a_i \mu_i E_P(f_i)\Big) p_I + \Big(\sum_{i=1}^{T-P+1} b_i \mu_i E_P(f_i)\Big) p_{II} \neq 0$$

for all choices $a_i = 0, 1$ and $b_i = -P, \ldots, P$. This directly follows from Equation 1.

In the case of unit weights and unit E_P's, the condition of uniqueness is that $ap_I + bp_{II} \neq 0$ for all non-negative integers a, b, such that $|a| \leq T - P + 1$ and $|b| \leq (T - P + 1) \cdot P$. In particular if p_I/p_{II} is irrational then the best weighted seasonal $[t_0, t_1]$-pull extrapolation is uniquely defined in such settings.

7. Detection of best weighted seasonal $[t_0, t_1]$-push extrapolation

Note that the push extrapolation loss functional $\Lambda_f^{t_0,t_1}$ is a piecewise linear function where $g(t_0), \ldots, g(t_1)$ are considered as variables. Global minima of such functions are obtained at points, where at least $t_1 - t_0 + 2$ different linear domains come together. This gives $t_1 - t_0 + 1$ linear equations on the variable values $g(t_0), \ldots, g(t_1)$.

Theorem 7.1. *There exists a minimizer (i.e., a best weighted seasonal $[t_0, t_1]$-push extrapolation) of $\mathcal{L}_f^{t_0,t_1,\mu}$ which is described as an intersection point of $t_1 - t_0 + 1$ planes of the following family:*

$$\sum_{t \in I} g(t) = \sum_{t \in I} f_i(t) + F_i^{t_1}(\tilde{t}) \tag{11}$$

where I is an arbitrary subset of the set $\{t_0, \ldots, t_1\}$, for the choice of normalized periodic particles we have $1 \leq i \leq T - P + 1$, and the integer value $\tilde{t}$ satisfies $t_1 - 1 \leq \tilde{t} \leq t_1 + P - 1$. (Recall that $F_i^{t_1}(t_1 - 1) = 0$.)

Proof. Since $\mathcal{L}_f^{t_0,t_1,\mu}$ is piecewise linear, its minimizer is at one if the vertices of $\mathcal{L}_f^{t_0,t_1,\mu}$, i.e., at intersection of $t_1 - t_0 + 2$ hyperplanes of nonlinearity of $\mathcal{L}_f^{t_0,t_1,\mu}$. Each of such hyperplanes is in fact a hyperplane of non-linearity for $\Lambda_{f_i}^{t_0,t_1}$ for some i and it is defined by one of Equations (11).

There are two type of hyperplanes where $\Lambda_f^{t_0,t_1}$ is nonlinear. The hyperplanes of the first type are defined by

$$R_{f,g}^{t_0}(t) = 0.$$

The hyperplanes of the second type are defined by

$$R_{f,g}^{t_0}(t) - F_i^{t_1}(\tilde{t}) = 0.$$

Similarly to the proof of Theorem 6.2, the time $\tilde{t}$ can be chosen from the set

$$\{t_1 - 1, \ldots, t_1 + P - 1\}.$$

Now both types are of the form of Equation (11). This concludes the proof. □

Acknowledgements. We are grateful to O. Gorelkina, C. Constantinescu-Loeffen, H. Assa, A. Lyons, S. Mitra, and K. Sharkey for useful remarks and comments. Oleg Karpenkov is partially supported by EPSRC grant EP/N014499/1 (LCMH).

References

[1] G.E.P. Box, G.M. Jenkins and G.C. Reinsel, *Time series analysis : Forecasting and control*, Pearson Education, Delhi, 1994.
[2] R.G. Brown, *Smoothing, Forecasting and Prediction*, Englewood Cliffs: Prentice Hall, 1963.
[3] R.L. Chambers, C.J. Skinner, (editors), *Analysis of Survey Data*, Wiley Series in Survey Methodology, Wiley, 2003.
[4] D.F. Findley, B.C. Monsell, M.C. Otto, W.R. Bell, M. Pugh *Towards X-12 ARIMA*, Technical report, Bureau of the Census, 1992.
[5] D.F. Findley, B.C. Monsell, W.R. Bell, M.C. Otto, B.C. Chen *New capabilities and methods of the X-12 ARIMA seasonal adjustment program*, Technical report, U.S. Bureau of the Census, 1996.
[6] R. Gençay, F. Selçuk, B.J. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, 2001.
[7] C.C. Holt, *Forecasting seasonals and trends by exponentially weighted moving averages*, ONR Research Memorandum 52, Carnagie Institute of Technology, Pittsburgh, Pennsylvania, 1957.
[8] A. Kolmogorov, *Interpolation and extrapolation von stationaren Zufalligen Folgen*, Bulletin of the Academy of Sciences, USSR, Ser. Math., 5, 3–14, 1941.
[9] N. Wiener, *The extrapolation, interpolation and smoothing of stationary time series with engineering applications*, Wiley: New York, 1949.
[10] P.R. Winters, Forecasting sales by exponentially weighted moving averages, *Management Science*, **1960**, *6*, 324-342.

E-mail address, Tetiana Boiko: `t.boiko@liverpool.ac.uk`
E-mail address, Oleg Karpenkov: `karpenk@liverpool.ac.uk`
E-mail address, Bulat Rakhimberdiev: `rakhimberdiev@mail.ru`